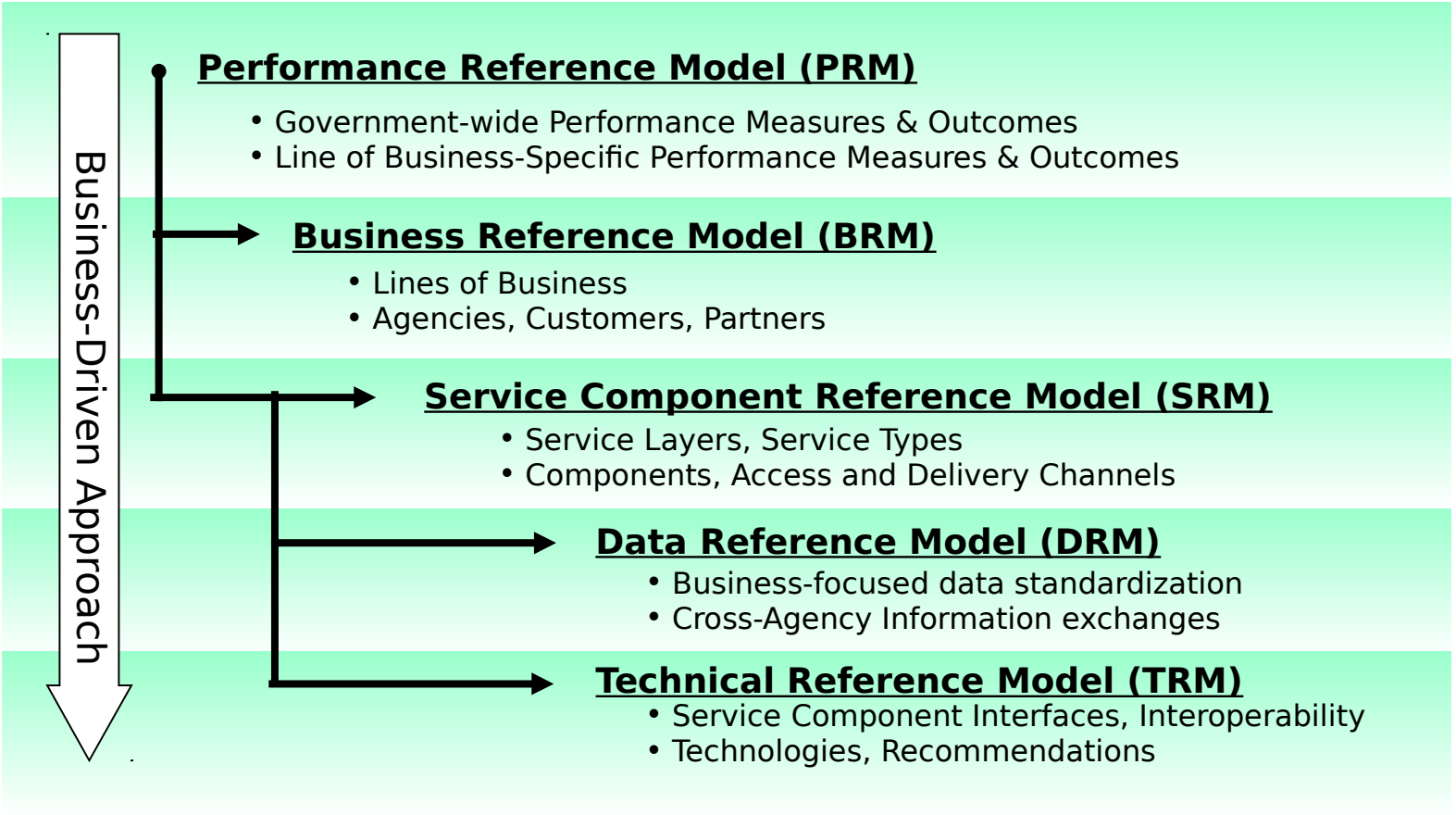
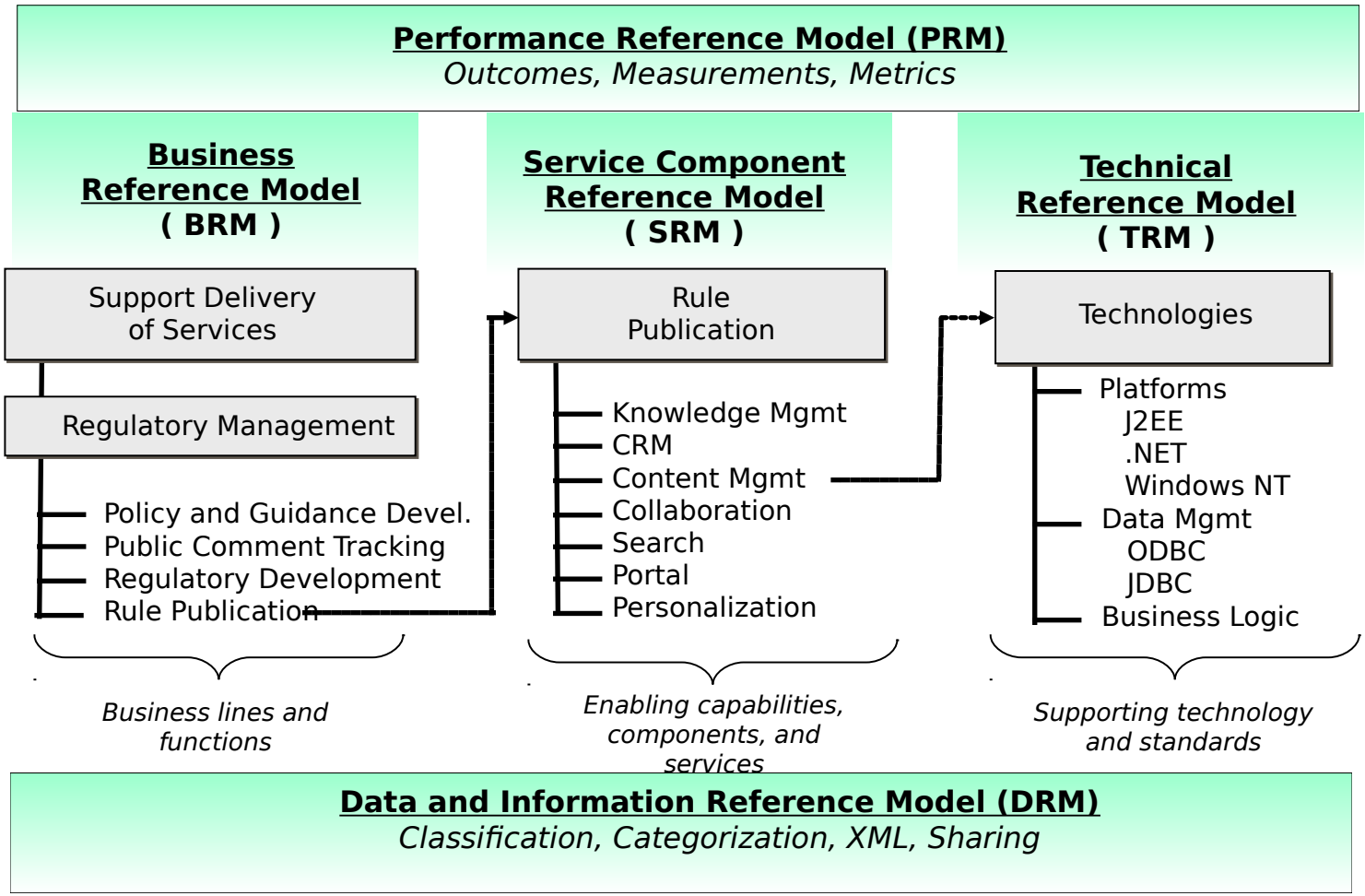
XML in the Emerging U.S. Federal Information Architecture

Presented by
Eliot Christian, USGS
April 30, 2003

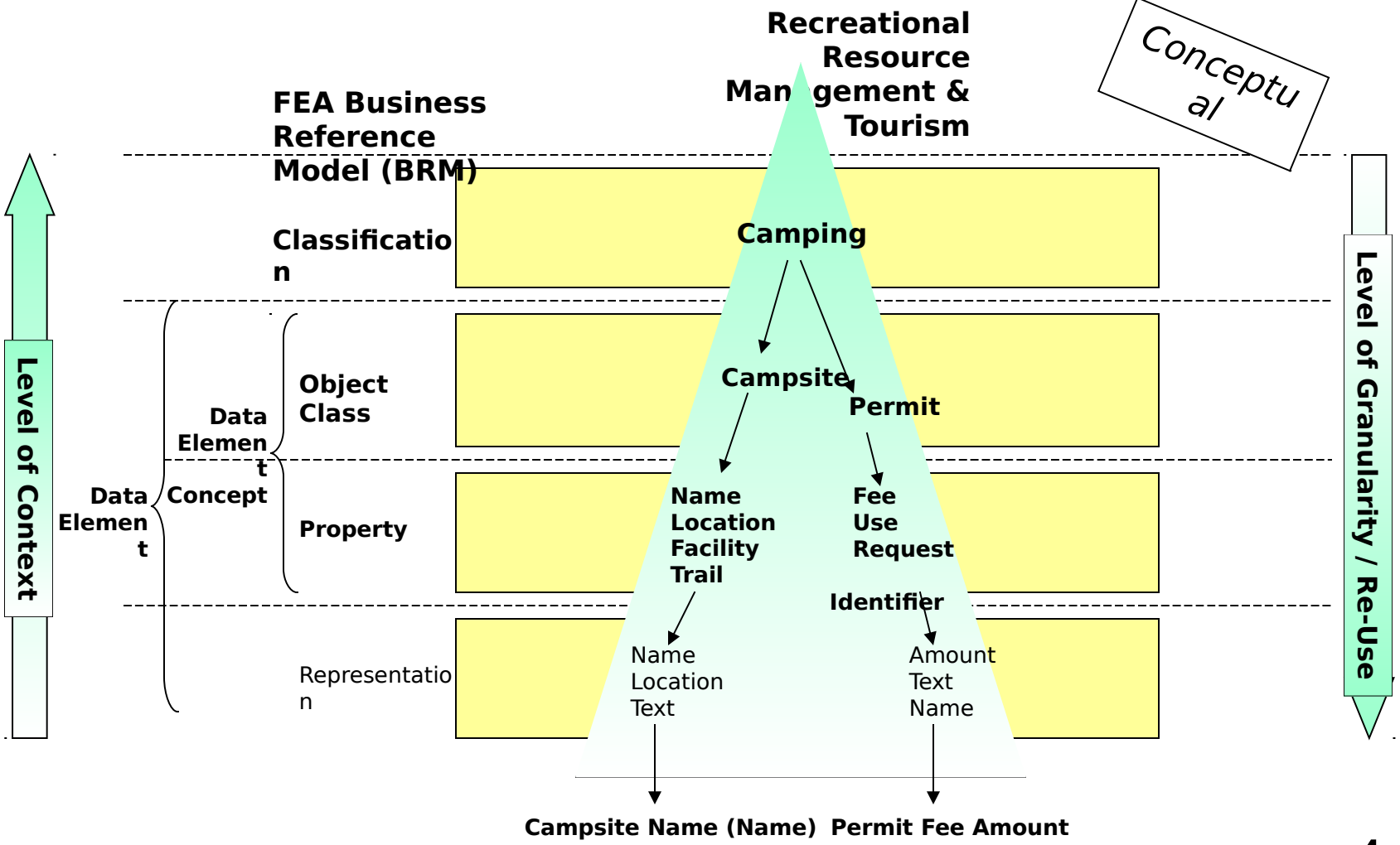
Federal Enterprise Architecture (FEA)


Business-Driven Approach
Performance Reference Model (PRM)
• Government-wide Performance Measures & Outcomes
• Line of Business-Specific Performance Measures & Outcomes
Business Reference Model (BRM)
• Lines of Business
• Agencies, Customers, Partners
Service Component Reference Model (SRM)
• Service Layers, Service Types
• Components, Access and Delivery Channels
Data Reference Model (DRM)
• Business-focused data standardization
• Cross-Agency Information exchanges
Technical Reference Model (TRM)
• Service Component Interfaces, Interoperability
• Technologies, Recommendations

FEA Reference Models



FEA Data Reference Model (DRM) Overview of Framework



Information Architecture Principles

1. Avoid non-standard data syntaxes
   Standard: XML or ASN.1
2. Register the semantics of shared data elements
   Standard: ISO 11179
3. Document service interfaces in a standard way
   Standard: CORBA IDL, ebXML, Web Services, UDDI
4. Implement standard for information discovery
   Standard: Information Retrieval, ISO 23950/GILS
5. Implement standard interfaces for geospatial data
   Standard: GILS/GEO, OpenGIS Consortium

Information Architecture Principle 1:

Avoid non-standard data syntaxes

- Need to agree on a small number of robust data representation syntaxes that handle all data types required to traverse the shared interfaces
- Agreement to use either ASN.1 (Abstract Syntax Notation) or XML (Extensible Markup Language
- Both are interconvertible and both are compatible with Web Services

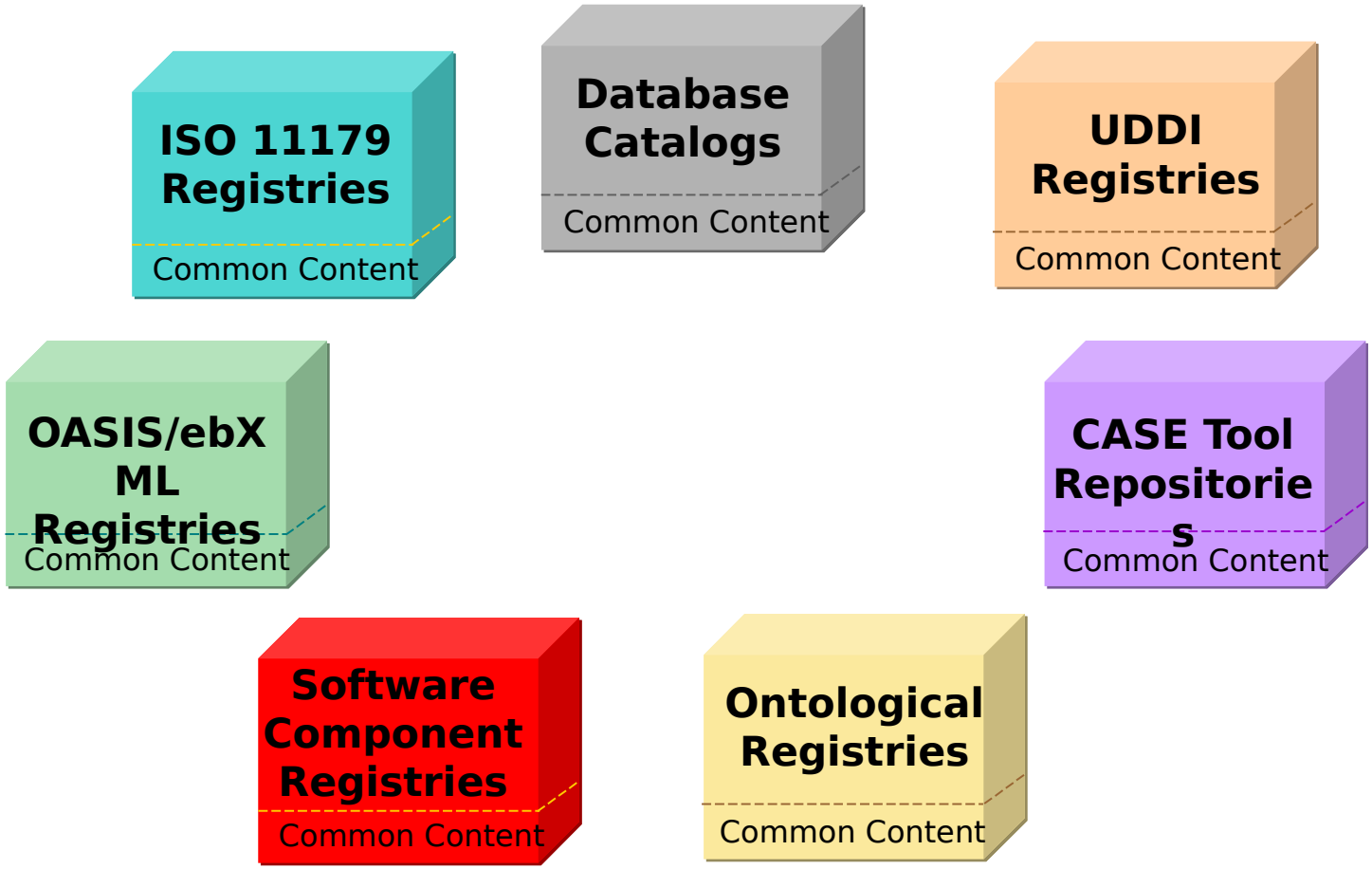
Information Architecture Principle 2: Register the semantics of shared data elements

- Need a common understanding of the meaning of data elements traversing the interface
- ISO 11179 (Information Technology--Metadata Registries) supports registration of data using virtually
  any syntax

Registries as Applied to Data/ Information/ Knowledge Management

- Several types of “registries” commonly used in data / information/ knowledge management
- Various registries have related content, extended and utilized in different ways
- The registries vary according to:
  - intended purpose
  - granularity of contents
  - the level of semantics management.

Types of Registries

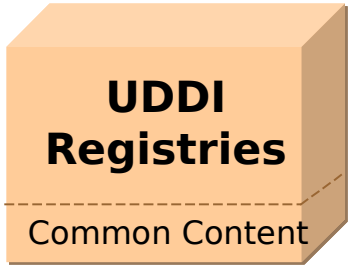


Types of Registries: Database Catalogs/Data Dictionaries/ Repositories

- Registers schema, integrity & operational information
- Includes all elements of database schemas: data elements, relations, integrity constraints
- Emphasis on information required to make database systems work for queries, etc., not semantic management

Types of Registries: Universal Description, Discovery, and Integration (UDDI) Registries

- Registers Web-based Business Services
- A White page directory of the business activities of firms
- Emphasis on interface specifications



Types of Registries: Case Tool (Encyclopedias/Repositories)

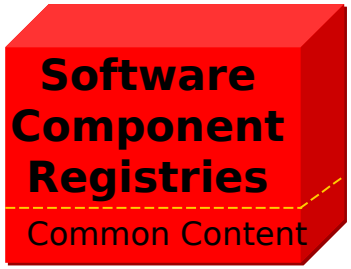
- Registers data models and application program logic
- Includes information needed to create a database and potentially the program code for a system, including database schemas
- Emphasis not on semantic management

Types of Registries: Ontological Registries

- Registers Concept Structures
- Includes relations among concepts (subsumption, inheritance,…) and axioms for inference among concepts, e.g., temporal/spatial reasoning, etc.
- Emphasis is on semantics

Types of Registries: Software Component Registries

- Registers reusable software components built on specific vendor platforms such as EJB, COM, CORBA IDL etc.
- Includes basic common elementary objects and object patterns
- Emphasis on reusability of software component to build flexible and standardized business systems.



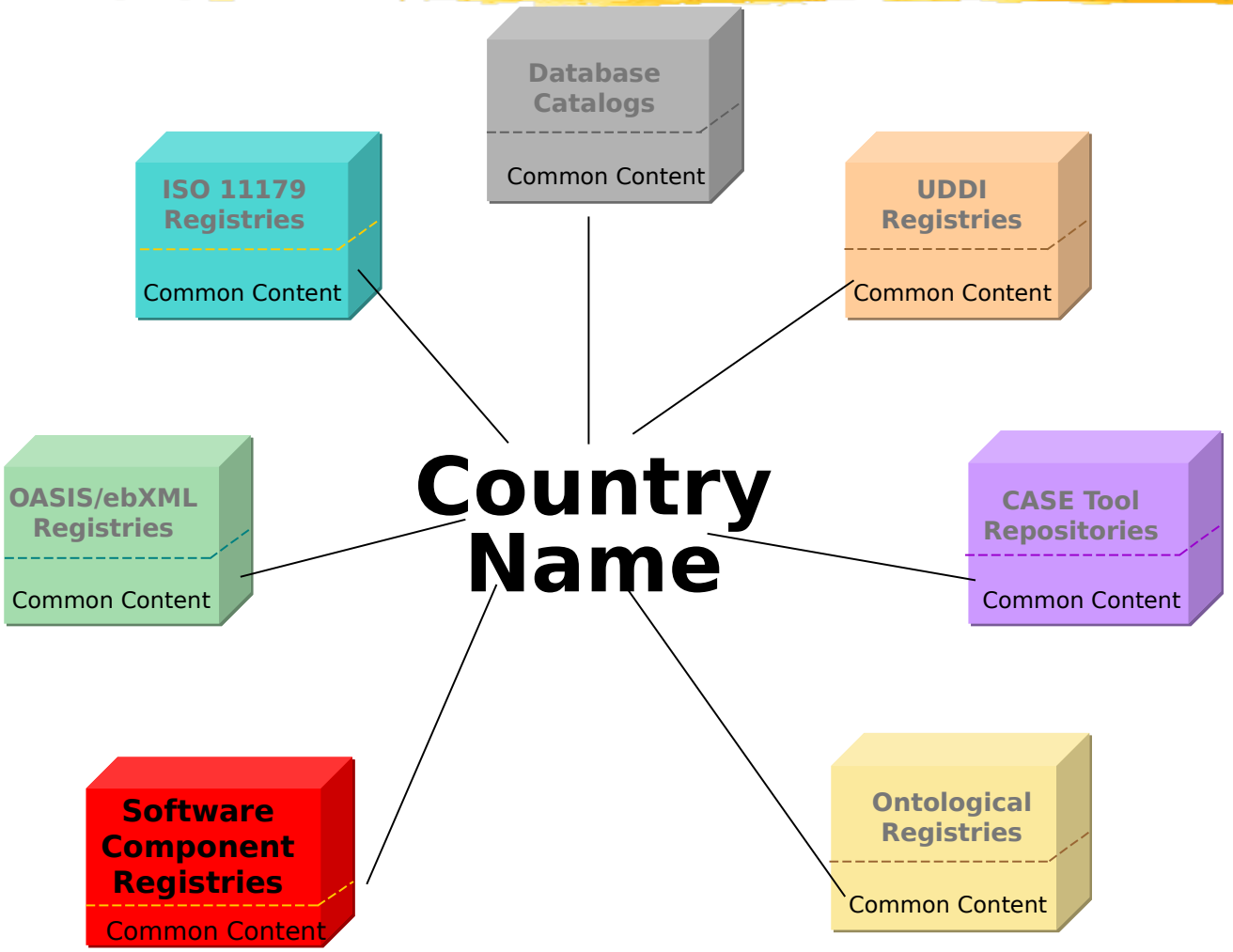
Types of Registries: OASIS/ebXML XML Registries

- Registers XML Artifacts
- Includes schemas and DTDs (XML schemata typically relate to a paper form, EDI document, or reporting requirement)
- Emphasis on syntactic (structure) information

Types of Registries:
ISO 11179 Metadata Registries

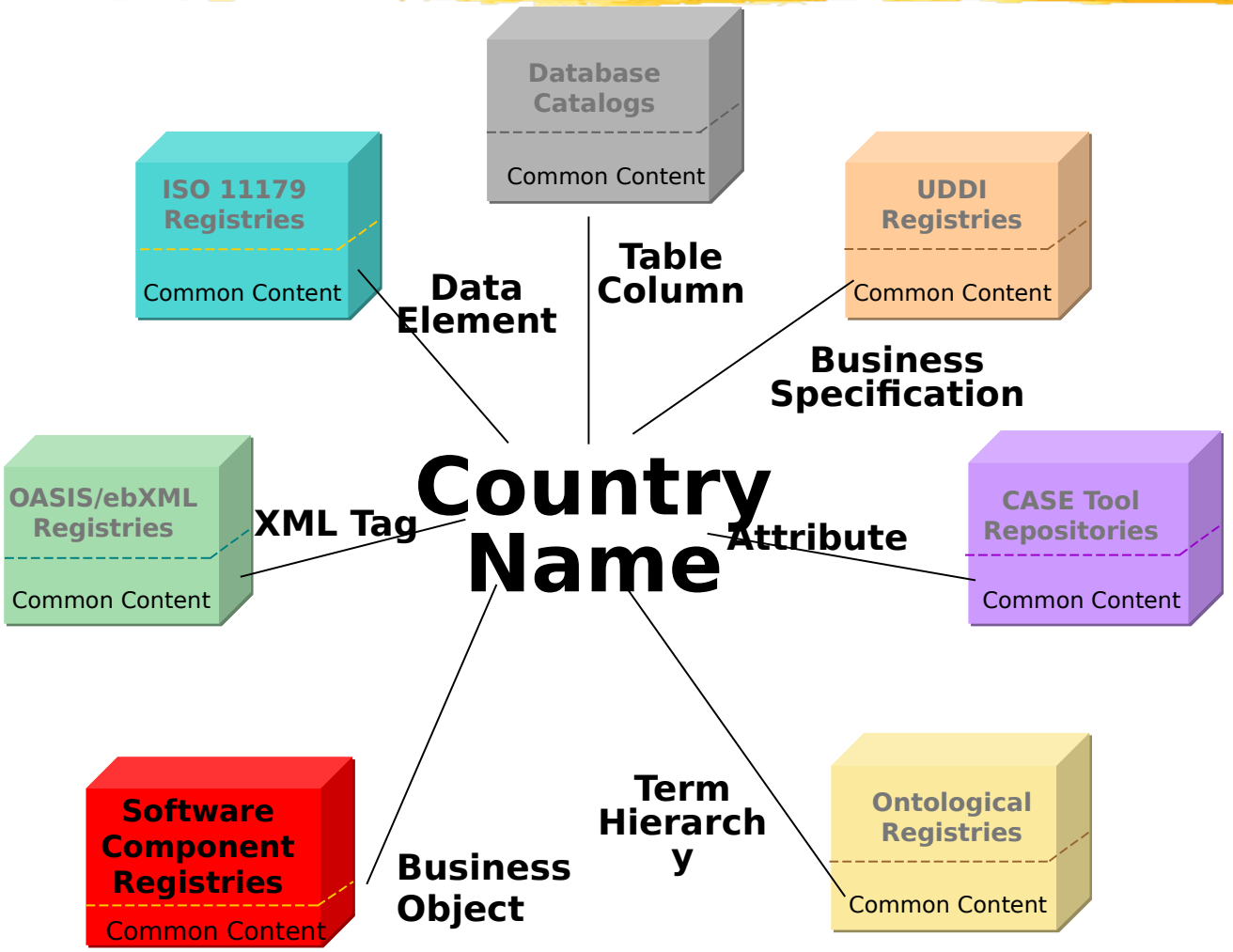
- Registers Data Elements, components of data elements and groups of data elements
  (e.g., country codes for customer place of residence)
- Includes data element concepts, data elements (including representation), value domains, and (multiple) taxonomies
- Emphasis on semantic information such as definitions of data elements and value meanings, and stewardship responsibilities

Example of Common Content across Registry Types

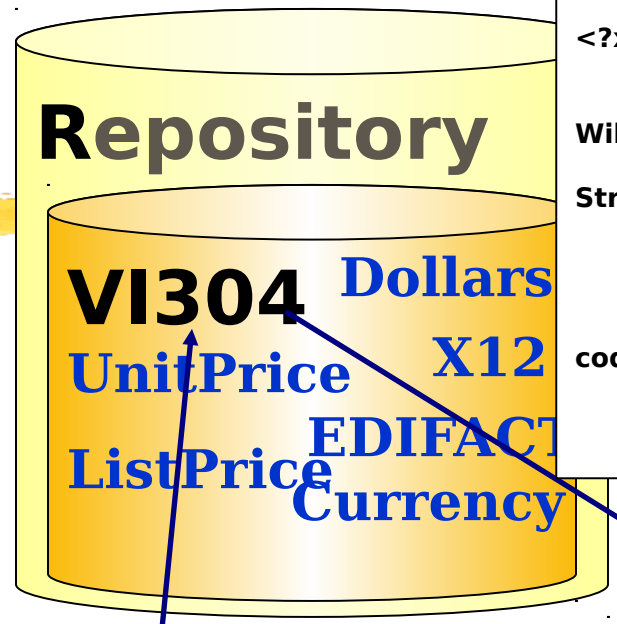


Registries

Example of Common Content



OASIS/ebXML XML Registry




```
<?xml version="1.0"?>
        <shipTo >
            <name>Alice
Wilson</name>
            <street>161 North
Street</street>
            <city>Happy Valley</city>
            <state>MO</state>
            <zip>63105</zip>
            <country
code>USA</country code>

        </shipTo>
```

Collaboration Partner #1

Collaboration Partner #2

Schema *or Template*

<Rep href= "http://www.disa.mil">DISA</Rep>

<ELEMENT rep= 'DISA' name='Country' Biz ='

Schema *or Template*

<Rep href= "http://www.disa.mil">DISA</Rep>

<ELEMENT rep= 'DISA' name='CntryCode' Biz

XML Instance

<Country>JPN</Country>

XML Instance

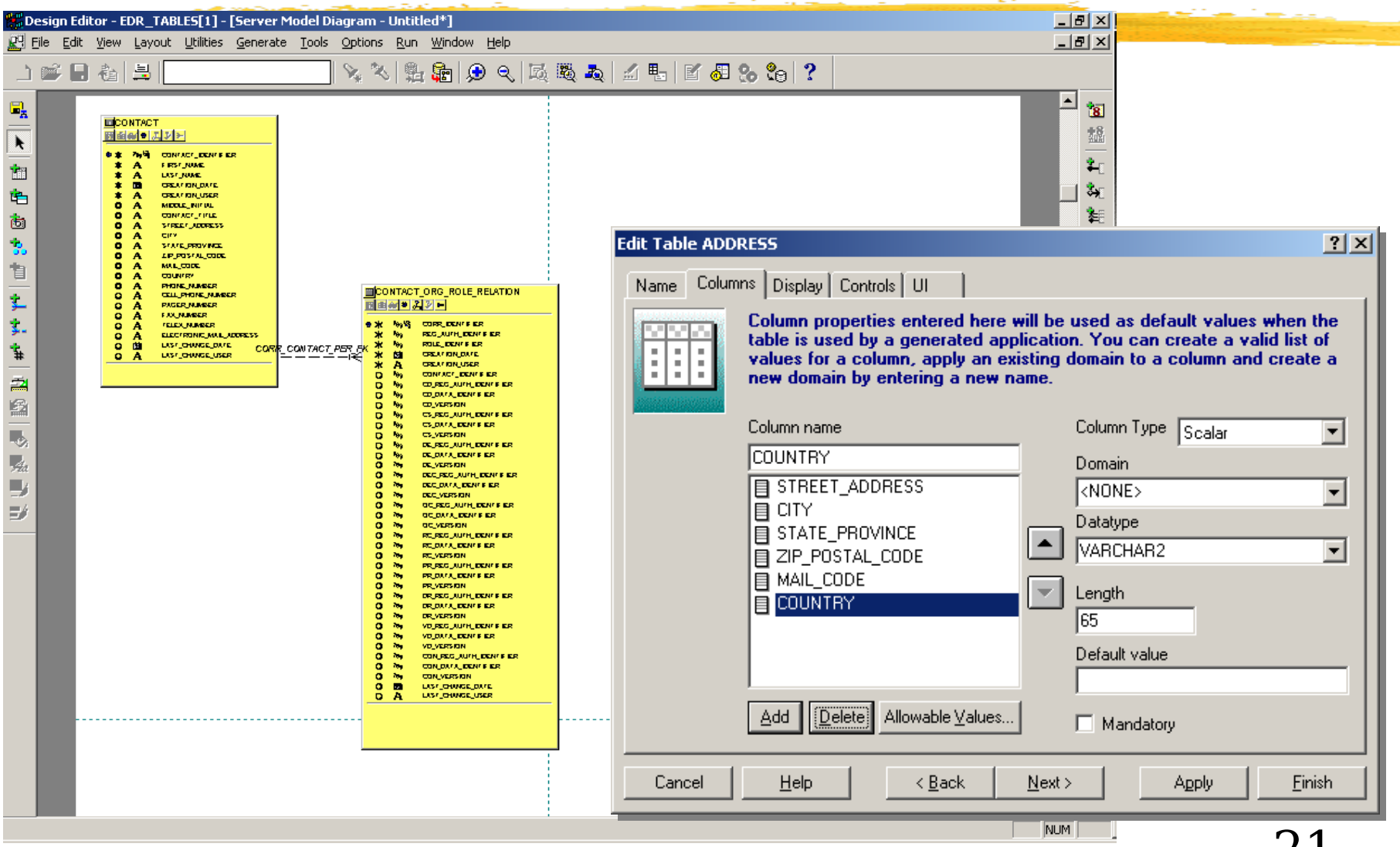
<CntryCode>JPN</CntryCode>

UDDI Registry

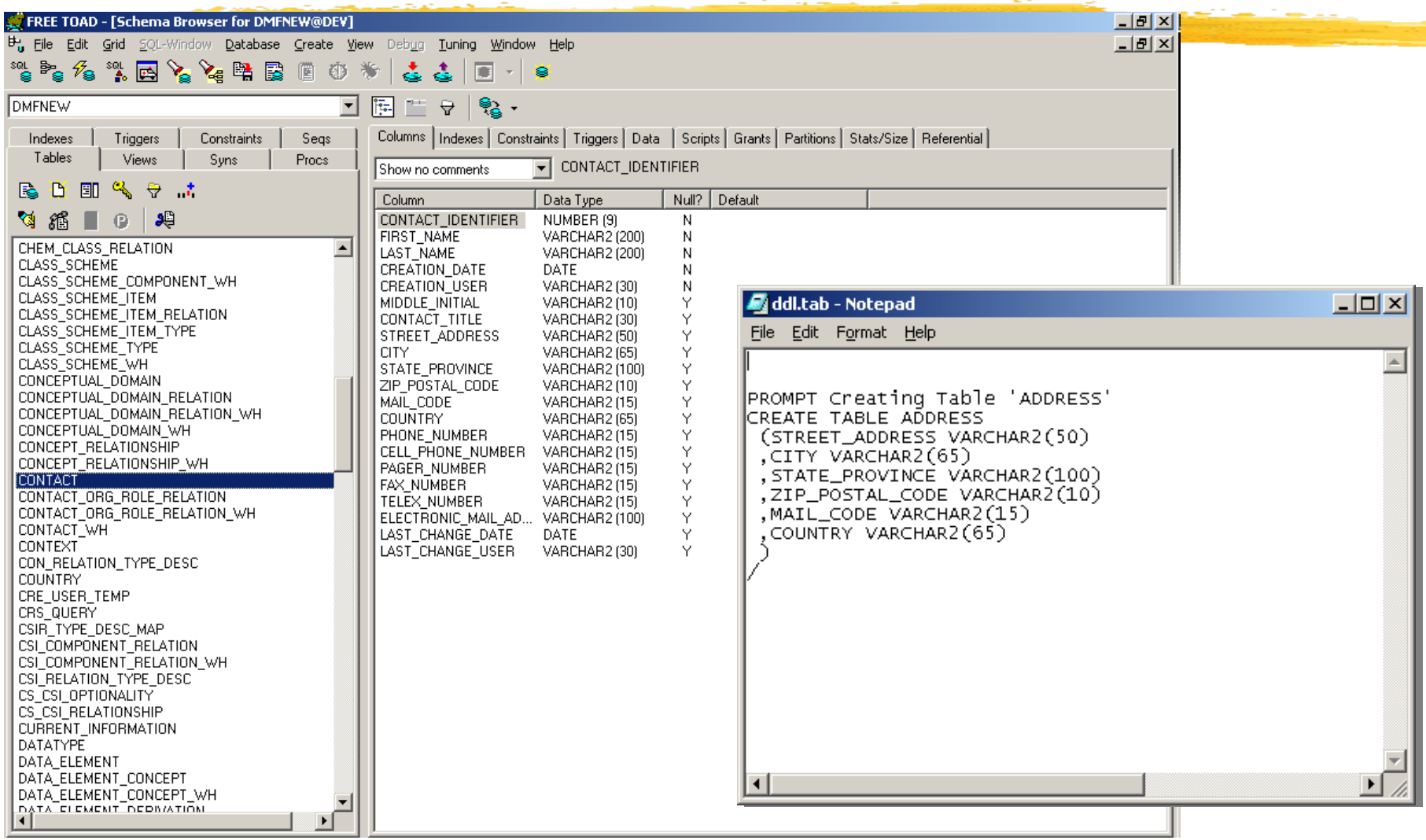
```
<?xml version="1.0"?>
            <shipTo>
               <name>Alice Wilson</name>
               <street>161 North
Street</street>
               <city>Happy Valley</city>
               <state>MO</state>
               <zip>63105</zip>
               <country code>USA</country
code>

            </shipTo>
```

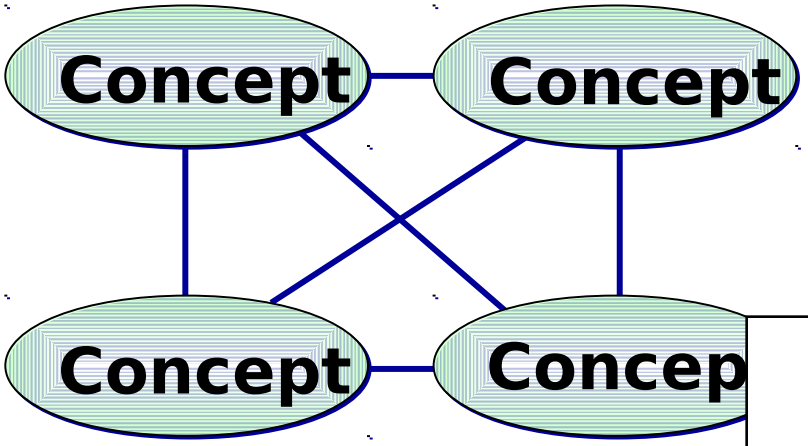
Case Tool Repository

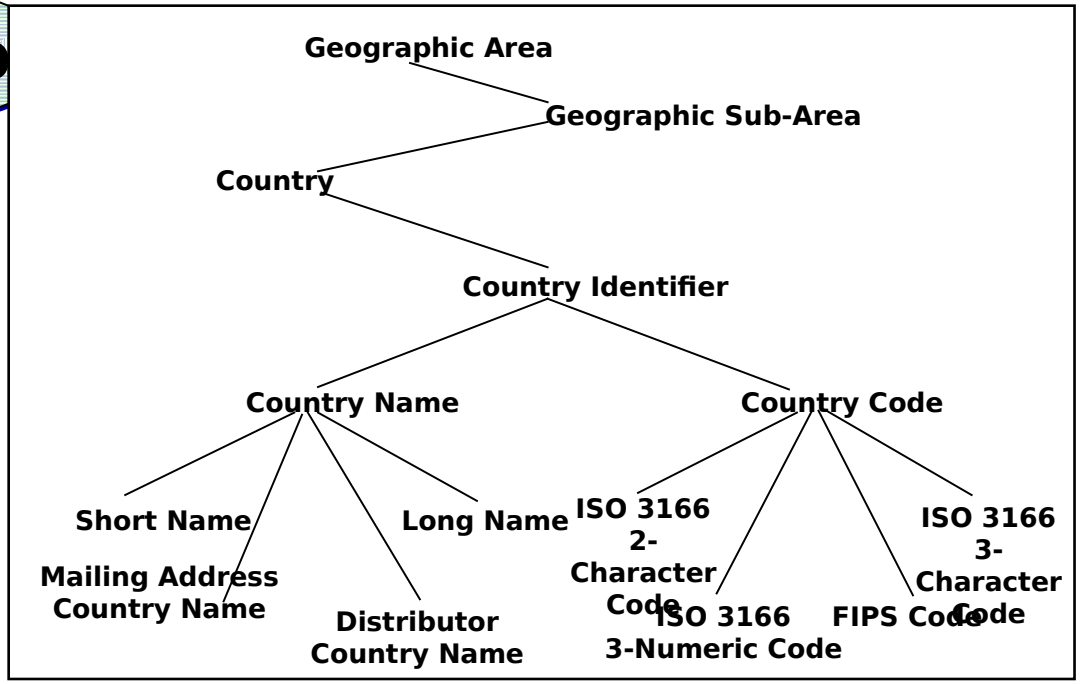


Database Catalog

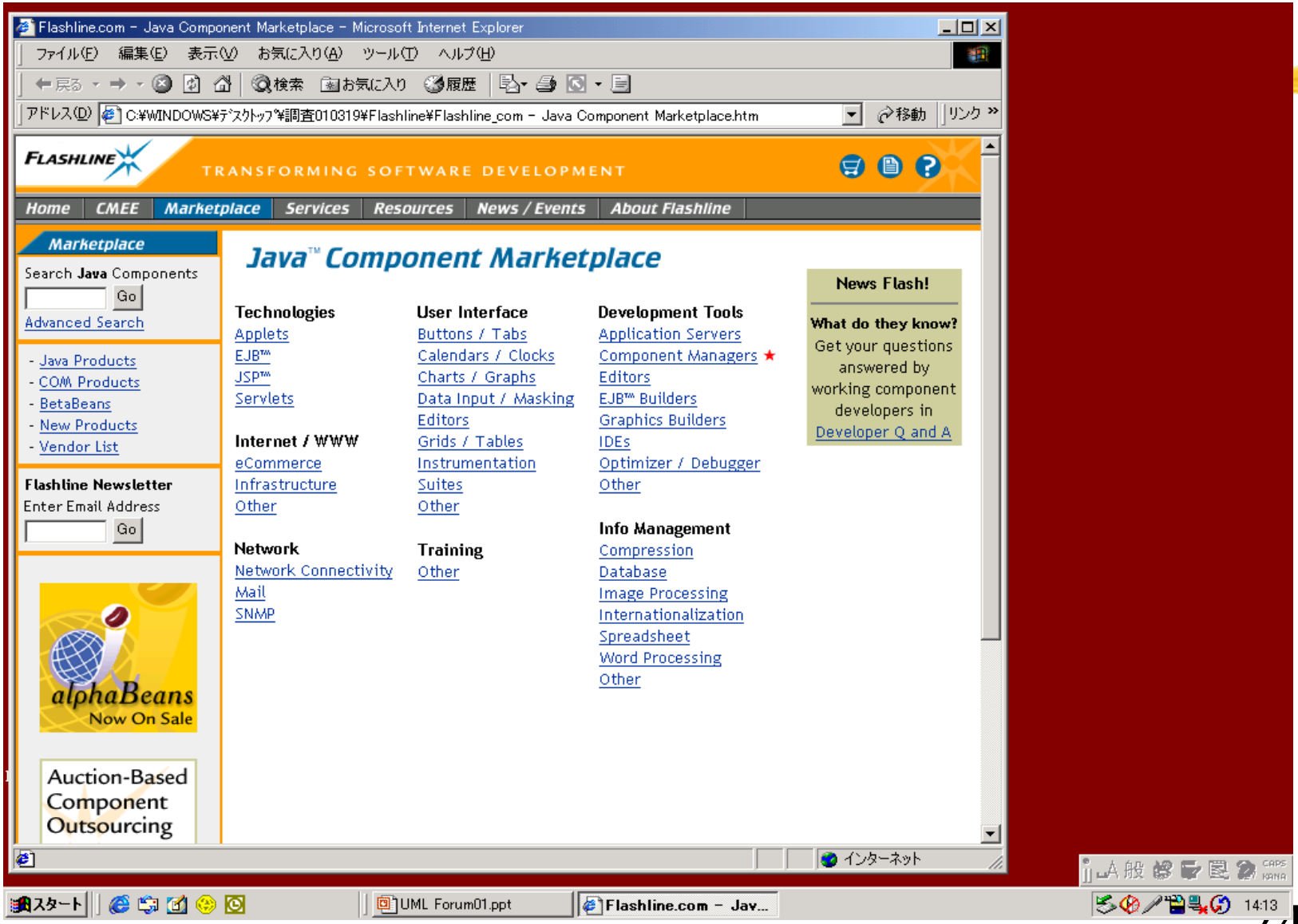


Ontological Registry

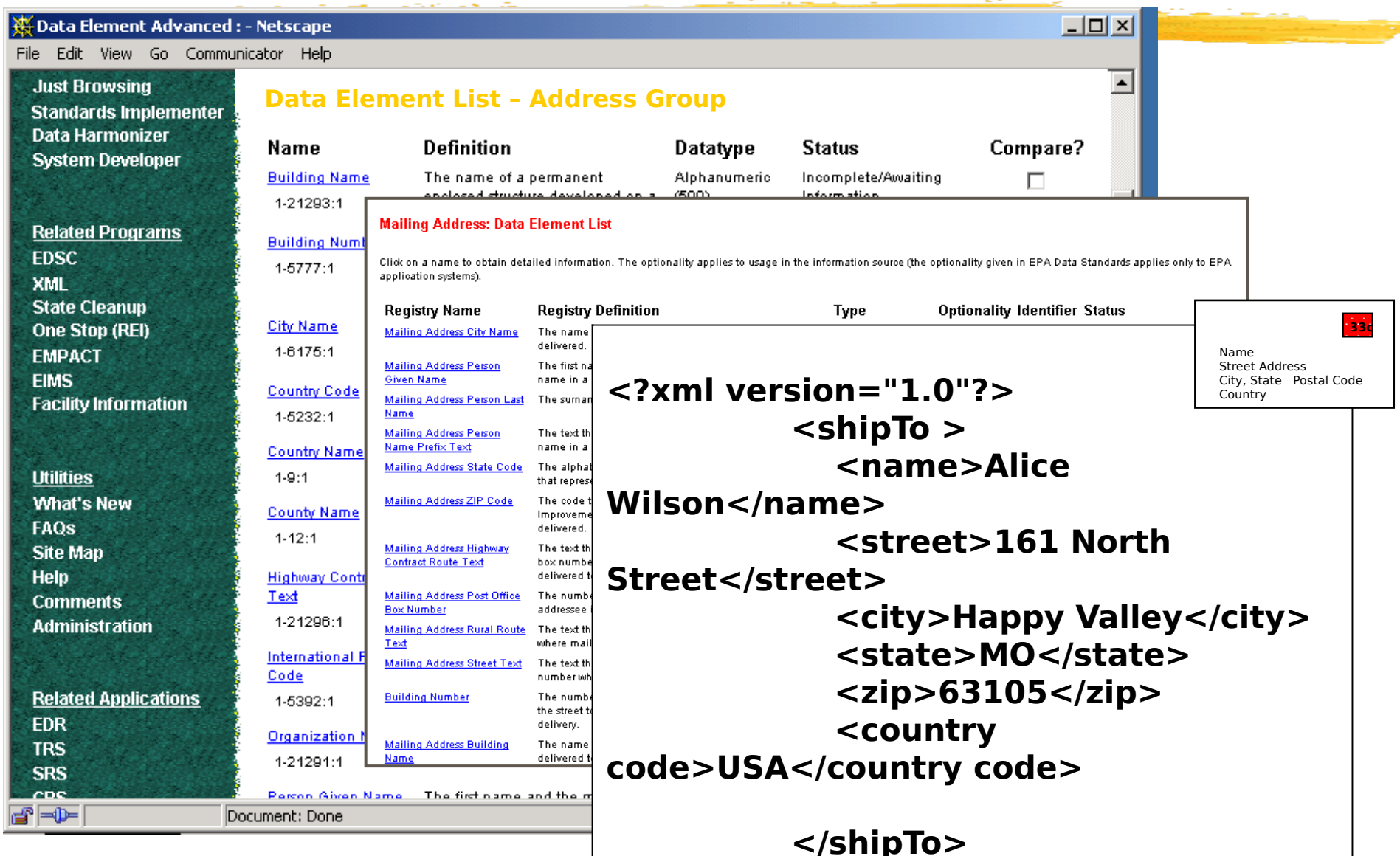




An Example of Software Component Registry



ISO 11179 Registry



Example of Common Content Country Identifier

Data Element Concept

Name: Country Identifiers
Context:
Definition:
Unique ID: 5769
Conceptual Domain:
Maintenance Org.:
Steward:
Classification:
Registration Authority:
Others

Algeria
Belgium
China
Denmark
Egypt
France
. . .
Zimbabwe

Data Elements

Name:
Context:
Definition:
Unique ID: 4572
Value Domain:
Maintenance Org.
Steward:
Classification:
Registration Authority:
Others

ISO 3166 English Name	ISO 3166 French Name	ISO 3166 2-Alpha Code	ISO 3166 3-Alpha Code	ISO 3166 3-Numeric Code
Algeria	L`Algérie	DZ	DZA	012
Belgium	Belgique	BE	BEL	056
China	Chine	CN	CHN	156
Denmark	Danemark	DK	DNK	208
Egypt	Egypte	EG	EGY	818
France	La France	FR	FRA	250
. . .	. . .	. . .	. . .	. . .
Zimbabwe	Zimbabwe	ZW	ZWE	716